AD_____

Award Number: DAMD17-98-1-8294

TITLE: Isolation of Genes from Chromosome Region Ip31 Involved in the Development of Breast Cancer

PRINCIPAL INVESTIGATOR: John K. Cowell, Ph.D.

CONTRACTING ORGANIZATION: The Cleveland Clinic Foundation
Cleveland, Ohio 44195

REPORT DATE: September 2000

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	September 2000	Annual (1 Sep 99 - 31 Aug 00)

4. TITLE AND SUBTITLE
Isolation of Genes from Chromosome Region Ip31 Involved in the Development of Breast Cancer

5. FUNDING NUMBERS
DAMD17-98-1-8294

6. AUTHOR(S)
John K. Cowell, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The Cleveland Clinic Foundation
Cleveland, Ohio 44195

E-MAIL:
cowellj@ccf.org

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT ***(Maximum 200 Words)***

The p31 region of chromosome 1 shows frequent (50%) loss of heterozygosity (LOH) in breast tumors. This observation implies the presence of a tumor suppressor gene in this region which, when mutated, contributes to tumorigenesis. The maximal common region undergoing LOH extends over approximately 2 Mbp. We have constructed a map of overlapping BAC clones spanning this region with only 4 gaps. The total available sequence from this region suggests that approximately 50-60% of it is now completed. Using gene analysis tools, we have been able to demonstrate that few full-length genes are located in this region and that the ESTs from the databases are clustered to a proximal position of the contig. The high level of sequencing completed from this region means that it will soon be possible to establish a baseline transcription map. The genes identified in this way can now be tested systematically for mutations in breast tumors. The progesterone receptor gene lies at the limit of the region suggesting that it was possibly the target for LOH. We established the exon/intron structure of this gene and undertook SSCP analysis of breast tumors. No gene-inactivating mutations were found excluding this gene from involvement in breast cancer development

14. SUBJECT TERMS	15. NUMBER OF PAGES
Breast Cancer	10
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

INTRODUCTION

One of the hallmarks of tumor suppressor genes is that both copies must be inactivated for the development of the malignant phenotype. A phenomenon associated with this inactivation is that a single recessive mutation is exposed following loss of a chromosomal region containing the normal copy of genes. This event is usually manifested in tumors as loss of heterozygosity (LOH) when compared to normal tissue from the same individual. Genome screening for LOH in breast cancer has identified chromosome region 1p31 as one commonly (50%) involved in LOH indicating the site of a tumor suppressor gene. By comparison of LOH in a large number of tumors, the minimal region has been defined by two markers which are 500-800 Kb apart. An extended region of loss spans approximately 2 Mbp. The overall aim of this project is to characterize the region of chromosome 1p31 which shows consistent loss of heterozygosity in breast tumors by constructing a BAC contig across the region, searching for genes in this contig and then performing a mutation analysis to determine which ones are involved in tumorigenesis. One gene, TTC4, was identified in this region and, after extensive analysis, could not be shown to be involved in breast cancer. It has since been shown that this gene is mutated and so potentially involved in the development of some melanomas. 1p31 is also frequently involved in LOH in melanomas. To identify more genes in the critical 1p31 region involved in breast cancer, we extended our search for candidates using a variety of approaches but, in particular, using the latest releases of sequence form the Human Genome Project.

BODY

In our previous report we described the construction of a YAC contig across the minimally deleted region which extended approximately 1500 Kb. In the early part of this year we began a series of hybrid cDNA capture experiments to isolate genes from a breast cancer cDNA library. This approach depends on hybridizing cDNA clones from a total library to YAC DNA and then, through successive rounds of amplification and selection, a number of clones are recovered for testing which apparently have homology to the target YAC clone. Because of the presence of repetitive sequences in many cDNA clones, however, the risk of isolating false positives is very high using this approach, especially with YACs since they are, on average, 1 Mbp long. After a series of experiments, we isolated a number of candidate cDNAs but were unfortunately not able to map any of the clones back to the YAC contig used in the screen. At this point, data from the Whitehead Institute describing new releases of the human gene map for chromosome 1, although not identifying any full-length genes in the critical 1p31 region, did map the progesterone receptor gene (PTGFR) to the distal border of this region. Because of the potential involvement of this receptor in breast cancer development, we decided to submit this gene to an analysis by SSCP/mutation screening and DNA sequencing. In order to establish the exon/intron structure, we identified a BAC clone which contained the PTGFR gene and then, by aligning the cDNA sequence with the genomic sequence, were easily able to establish the exon structure. The PR gene only contains three exons and so primers were designed (Table 1) to amplify the individual exons and subject them to SSCP analysis. This approach identifies single strand PCR products which show altered mobility on acrylamide gels, indicative of a DNA sequence change (either mutation or polymorphism). Since each of the PTGFR exons were large (>250 bp), we designed a series of overlapping PCR products to cover each exon (Table 1). We used 32 tumor samples from breast cancers in this analysis and identified a number of band changes. When sequenced, these changes proved to be silent polymorphic variants resulting

in silent changes in the amino acid code for the gene (Table 2). From this survey, therefore, we concluded that PR gene is not frequently mutated in breast cancer.

At this point the accelerated pace of the human genome sequencing project dictated an alternative approach, which was to build up BAC contigs from sequence databases across the 1p31 region and then use database searching to identify ESTs and genes. Using the available physical map which was used to created the YAC contig, we searched the high-throughput genome sequencing database for BACs across the region. These are shown in Figure 2. These individual BACs were identified and then overlapping BACs from the same databases were identified by BLAST searches. The sequencing effort is being undertaken by the Sanger Center in the UK and the genome center at Washington University. Since there is considerable redundancy in these two efforts, it was possible to build up a contig of BACs with only four gaps, the lengths of which cannot be determined at this stage. Of the 2-3 Mbp region, approximately 50% has been sequenced; thus, there is adequate sequence to begin to build up a transcription map of the region to search for candidate genes.

Using BLAST searches and the Genome Channel program (GenScan) at Oak Ridge National Laboratories, we were able to identify ESTs from the region. These are summarized in Table 3. Firstly, It has been an important observation that there are few genes for which the full-length sequence to map this region is known. A number of ESTs were identified, but most of these cluster at the proximal end of the contig and include the latrophilin gene which has previously been shown not to be mutated in breast cancer cells. The only other gene in this region is the DRIL2P gene which is a pseudogene for the DRIL transcription factor. The distal 60% of this contig apparently contains no ESTs, either through BLAST searches or through analysis of the sequence from the individual BACs using GenScan. Although a number of individual exons are predicted in this region using GRAIL, no extensive gene structures could be identified. Even though these gene prediction tools are notoriously inaccurate in their ability to identify genes, it is still true that the region is very gene poor which has the advantage that any gene which is found has a higher probability of being the one reduced to homozygosity in breast cancers.

Within the proximal clusters of ESTs, a number show homologies to Unigene clusters which provide for more cDNA sequence from these genes. Despite this, the available sequence still does not provide any indication of the potential function of any of these genes. We are now concentrating on using the available sequence information to screen cDNA libraries to identify more full-length clones.

KEY RESEARCH ACCOMPLISHMENTS

- Construction of a BAC contig across the LOH region in 1p31.
- Analysis of approximately 1700 kb of sequence using GenScan to predict genes in the critical region.
- Identification of ESTs and Unigene clusters in the 1p31 region
- Exclusion of the progesterone receptor gene in the development of breast cancer

REPORTABLE OUTCOMES

None

CONCLUSIONS

As a result of advances in the development of the human genome sequencing project, we have abandoned traditional gene cloning strategies and have concentrated on building up the contig of complete sequence data for the critical BRCA region in 1p31. This analysis

identifies approximately 1.7 Mbp of the region with only 4 gaps. An analysis of this data places the progesterone receptor at the outer limit of the contig and mutation analysis of this gene does not support its role in the consequences of LOH . We have identified 27 ESTs to date in the critical region and five Unigene clusters. None of the data available for genes in this region allows for a prediction of their function. The construction of a rudimentary transcription map for the 1p31 region involved in breast cancer now means that the isolation of full length genes will be possible by cDNA library screening. As the sequence data become more complete, the generation of a full transcription map will become possible.

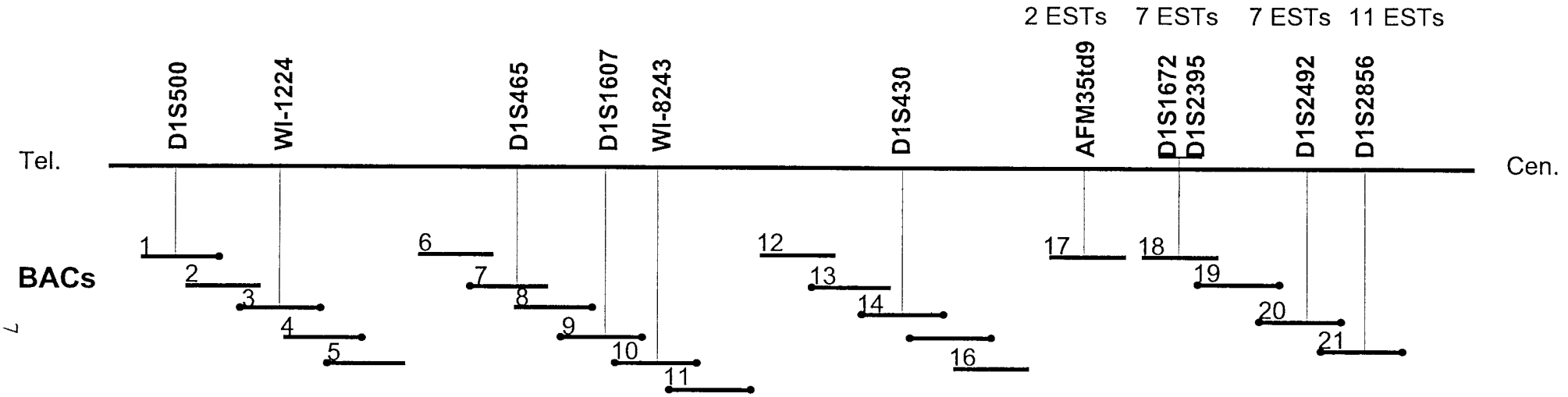
REFERENCES

None

APPENDICES

None

Figure 1: Summary of the BAC map of the BRCA 1p31 region relative to the physical map.
The location of ESTs identified by BLAST searches are shown.



1: RP11-239A21, Acc# AC025995
2: RP11-456F13, Acc# AL358533
3: RP11-316O2, Acc# AC073086
4: RP11-483I8, Acc# AL360080
5: RP11-466L22, Acc# AL355876

6: RP11-3B10, Acc# AC024634
7: RP5-1159G19, Acc# AL356501
8: RP5-887A10, Acc# AL391062
9: RP5-933B4, Acc# AL136234
10: RP11-115A15, Acc# AC009936
11: RP11-339D23, Acc# AC024634

12: RP11-82L20, Acc# AL354949
13: RP11-316C12, Acc# AC068047
14: RP1-54E18, Acc# AL356856
15: RP11-115M14, Acc# AC073627
16: RP11-417C2, Acc# AC027532

17: RP3-445O10, Acc# AL158173

18: RP5-837I24, Acc# AL138799
19: RP4-653E17, Acc# AL138841
20: RP11-100D1, Acc# AL358939
21: RP11-211B4, Acc# AL359705

Table1: Primers sequences, PCR product size and PCR conditions used for SSCP analysis for PTGFR gene.

Primer name	Sequence	PCR product Size	Annealing T°C
PTGFREX1F PTGFREX1R	AAC GAG TGC GCG GAG GGG AC CTG GTC CAA ACT CAG TGT CAG G	300bp	65
PTGFREX2-1F PTGFREX2-1R	GGC CAG ATA AAA CCC ATC CC AAG ATT CCC ACT GTC ATG AAG A	302bp	57
PTGFREX2-2F PTGFREX2-2R	CCT GCA GCT GCG CTT CTT TC CCA TGC AGA TAC CAA AAA TAC TG	313bp	60
PTGFREX2-31F PTGFREX2-3R	CTC ATC AAT GGA GCC ATA GC CGC CTG AAT TTT ATA GTC TCG	300bp	53
PTGFREX2-4F PTGFREX2-4R	CGA AAA TTA CAT CCA AAC AT GAG CTG GAT TAC CAT TTC CA	320bp	53
PTGFREX2-5F PTGFREX2-5R	ATG CAA TCA CAG GAA TTA CAC CTG AGT GCT GAA ATA AAT TTT G	265bp	55
PTGFREX3-1F PTGFREX3-1R	AAA CAA TAG CAT CAC TCT GTG CAT GCA CTC CAC AGC ATT GA	300bp	60
PTGFREX3-2F PTGFREX3-2R	CCGAAT GGC AAC ATG GAA TC AGG TAT TTA ACT AAC TGA AAT ATTG	290bp	55

Exon-Intron Boundries for PTGFR gene were determined by aligning the cDNA sequence of PTGFR with the unfinished sequence of BAC clone RP5-944H6 (Acc# AL136324).

Table 2: Summary of the silent polymorphisms found in the PTGFR gene in breast cancers. The 2 polymorphisms with a minus (-) annotation occur within the intron for exon 2.

Exon	nucleotide	amino acids
2	C63T	T21T
2	-152 T -> C	
2	–62 T -> C	
2	G435A	T145T

BAC clone Acc#	EST Acc#	UniGene Cluster	Gene name
RP3-445O10	AA909664	Hs.132428	
AL158173	AA827584		
RP5-837I24	R10940		
AL138799	AI147471	Hs.146741	
	AA406063	Hs.98003	
	BE894577		
	BE279803, BE256076		DRIL2P
	AI267423		
RP4-653E17	AI732798, AA487326	Hs.188839	
AL138841	AW105209	Hs.242795	
RP11-100D1	AW270271, AI243907	Hs.149086	
AL358939	AA347902		
	W03266		
	AA347901		
	N69441		
RP11-211B4	AI005162		
AL359705	AA369954		
	R01807		
	AA347901		
	AA259171		
	BE066265		
	T06662		
	H63565		
	H63611		
	BE709816		
	AL040055	Hs.24212	LATROPHILIN

Table 3: List of EST markers and their respective UniGene clusters that map to BAC clones from the breast cancer critical region in 1p32.